

Is GVB-CI Superior to CASSCF?

JOHN CULLEN

Department of Chemistry, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada

Received 11 September 1998; accepted 11 February 1999

ABSTRACT: Presently the most reliable approach for the study of reaction pathways where chemical bonds are broken and formed is to carry out CASSCF calculations followed by corresponding multireference perturbation or CI treatments. The latter step generally relaxes the “antibonding character” of the CASSCF results. In this study we demonstrate that similar results can be well approximated by using a less optimized MCSCF method and not performing the multireference perturbation or CI step at all. This is accomplished by performing a complete CI calculation within the active orbital space of the generalized valence bond perfect pairing (GVB-PP) model. The local bond/antibond character of the orbital space of the GVB-PP method also allows development of a fast, but robust, Bethe–Goldstone algorithm, which reconstructs the CI energy to an accuracy of a few tenths of a millihartree for most types of bond breaking cases found in chemical reactions. This algorithm executes at a speed proportional to N_p^4 where N_p is the number of localized electron pairs in the active space. © 1999 John Wiley & Sons, Inc. *J Comput Chem* 20: 999–1008, 1999

Keywords: GVB; CI; Bethe–Goldstone; CASSCF; *ab initio*; nondynamic electron correlation

Introduction

One of the most widely used *ab initio* procedures currently available for studying energy surfaces of chemical reactions is the complete active-space self-consistent-field (CASSCF) method. From an initial reference function a full

configuration interaction (CI) treatment is carried out over a small preselected subset of valence orbitals that is simultaneously optimized along with all other occupied molecular orbitals to self-consistency. Because the number of excited configurations that must be considered grows factorially with the size of the set of valence orbitals, known as the active space, and the number of electrons, practical calculations rarely exceed 10 active electrons and 10 active orbitals [1]. In addition, CASSCF calculations are generally not quantitative

Correspondence to: J. Cullen; e-mail: cullen@cc.umanitoba.ca

in accuracy, except in rare cases such as in the BH and C₂ molecules.^{2,3} The traditional approach to correct these deficiencies has been to carry out a second step, which uses multireference CI or perturbation theory to take into account the so-called dynamic electron correlation energy from excited configurations that lie outside the active space.⁴ Generally, CASSCF solutions are found to have too much antibonding character and, at equilibrium geometries, bond lengths are characteristically too long, while force constants are underestimated.⁵ In contrast, Hartree–Fock solutions usually exhibit too much bonding character, resulting in the opposite effects; that is, underestimations of bond lengths and overestimations of force constants.⁶ This raises the question whether a less optimized multiconfigurational SCF method positioned approximately midway between Hartree–Fock and CASSCF procedures could balance out these systematic errors. One possible choice that might accomplish this is the generalized valence bond (GVB) CI method.⁷ This method involves performing a full CI over the active space of localized bonding and antibonding orbitals generated by the GVB perfect pairing algorithm. The GVB-PP algorithm ignores all interbond electron correlation focusing exclusively on intrabond electron correlation producing an optimized local set of bond/antibond orbitals for each electron pair in a chemical bond. Evidence that such a procedure is a balanced approach can be seen in the excellent agreement with experiment Harding and Goddard obtained long ago for the excited-state spectra for small organics such as formaldehyde.⁸ Also, using a restricted CI in which only the dominant localized excitations are included. Carter and Goddard have successfully predicted the singlet–triplet gaps in carbenes.⁹

In this article we further examine the GVB-CI approach for molecular ground states at equilibrium geometries as well as bond-breaking configurations comparing these results with Hartree–Fock, CASSCF, full CI, and other high-level electron correlation methods. In addition, the localized nature of the bond/antibond orbital active space is used to piece together an approximate full CI treatment from variational solutions of clusters of interacting electron pair molecular fragments. The procedure, which executes in a computational step proportional to $N_p^3 - N_p^4$, where N_p is the number of electron pairs, provides results for most cases within a few tenths of a millihartree of the exact full CI results at a fraction of the computational

cost of exact full CI results. Calculations are presented for a variety of basis sets ranging from minimum STO-3G used to check the limitations of this algorithm to large correlation-consistent polarized-valence quadruple-zeta (CC-pVQZ) sets used for comparison with experiment. Semiempirical results are also reported. Before describing this algorithm as well as presenting an analysis of our results we begin with a brief review of the GVB-PP model.

GVB-PP Model

Originally developed in its strictly variational form by Goddard,¹⁰ the GVB-PP wave function is represented by an antisymmetrized product of doubly occupied orbitals, ϕ_i , singly occupied orbitals of same spin, ξ_j , and two electron functions or geminals, g_k . Each g_k is constructed from local ground and doubly excited configurations of a pair of optimized bonding and antibonding orbitals, ϕ_k, ϕ_k^* , respectively. When the strong orthogonality condition is imposed:

$$\langle g_k | g_l \rangle = 0, \quad k \neq l$$

GVB-PP becomes equivalent to the antisymmetric product of the strongly orthogonal geminal (APSG) method of Hurley, Lennard–Jones, and Pople.¹¹ Recently, we showed that, upon renormalization, the GVB-PP wave function can be represented as a constrained coupled cluster type:

$$|\Psi\rangle = e^{T_{\text{GVB}}} |0\rangle \quad (1)$$

where $|0\rangle$ is the single reference state of core orbitals and singly or doubly occupied bonding orbitals. The corresponding coupled cluster operator, T_{GVB} , is defined by:

$$T_{\text{GVB}} = \sum_k t_k a_k^{*\dagger} a_k^{-* \dagger} a_k^- a_k \quad (2)$$

where the summation is over the set of GVB-PP geminals. The creation operators for electrons in antibonding orbitals ϕ_k^*, ϕ_k^{-*} are denoted by $a_k^{*\dagger}, a_k^{-* \dagger}$, respectively, and a_k, a_k^- represent annihilation operators for electrons in bonding orbitals, ϕ_k, ϕ_k^- . The coefficients t_k can be found variationally, or using the coupled cluster ansatz. Numerically, results are almost identical. However, the couple cluster solution is analytical for the energy

and is given by:

$$E = E_0 + \frac{1}{2} \sum_k \left[b_k + 2a_k - (b_k^2 + 4a_k b_k + 8a_k^2)^{1/2} \right] \quad (3a)$$

$$E_0 = \sum_k n_k \langle k|h|k \rangle + \frac{1}{2} \sum_k \sum_l n_k n_l \left[(kk|ll) - \frac{1}{2} (kl|kl) \right] \quad (3b)$$

$$a_k = (k^*k|k^*k) \quad (3c)$$

$$b_k = 2(\varepsilon_k^* - \varepsilon_k) - 4(k^*k^*|kk) + (kk|kk) + (k^*k^*|k^*k^*) \quad (3d)$$

ε_k^* , ε_k are the diagonal elements of the Fock matrix f , defined by:

$$f_{kl} = \langle k|h|l \rangle + \sum_r n_r \left[(kl|rr) - \frac{1}{2} (kr|lr) \right] \quad (3e)$$

Here, n_r is the orbital occupancy for the reference $|0\rangle$.

The remaining problem of finding optimized orbital solutions is solved by direct minimization of the energy using the BFGS variable metric algorithm coupled with the Fischer–Almlof¹³ technique, which linearizes the storage requirements for the inverse Hessian. This procedure was used in our calculations to generate the valence bond/antibond orbital space. Further details of the method can be found in ref. 12.

Localized Cluster CI Algorithm

A recurrent theme for treating large molecular systems is the representation of a molecule as an ensemble of interacting localized electron pairs.^{14–17} Within such a framework, energies $E(i)$, $E(ij)$, $E(ijk)$, $E(ijkl)$, for groups of one, two, three, and four localized electron pairs i , j , k , l , respectively, can be computed separately with local subsets of orbitals. A perturbation series can then be constructed for the total energy as:

$$E = E_0 + \Delta E \quad (4a)$$

$$\Delta E = \sum_i E(i) + \sum_{i>j} \Delta E(ij) + \sum_{i>j>k} \Delta E(ijk) + \sum_{i>j>k>l} \Delta E(ijkl) + \dots \quad (4b)$$

The two-pair, three-pair, four-pair, etc., interaction energies $\Delta E(ij)$, $\Delta E(ijk)$, $\Delta E(ijkl)$ are defined by:

$$\Delta E(ij) = E(ij) - E(i) - E(j) \quad (4c)$$

$$\Delta E(ijk) = E(ijk) - \Delta E(ij) - \Delta E(ik) - \Delta E(kl) - E(i) - E(j) - E(k) \quad (4d)$$

$$\Delta E(ijkl) = E(ijkl) - \Delta E(ijk) - \Delta E(ijl) - \Delta E(ikl) - \Delta E(jkl) - \Delta E(ij) - \Delta E(ik) - \Delta E(il) - \Delta E(jk) - \Delta E(jl) - \Delta E(kl) - E(i) - E(j) - E(k) - E(l) \quad (4f)$$

At the post-Hartree–Fock level, this simple procedure was originally used by Nesbet¹⁴ to compute electron correlation energies of all second-row atoms with a purported accuracy ranging from 98.5% to 100.3% of the empirical correlation energy. Physically, successive variational solutions are found for effective Schrödinger equations of clusters containing one, two, three, etc., electrons embedded in the Fermi sea of the remaining electrons of an N -fermion system. The resulting equations are a generalization of a pair correlation method derived by Bethe and Goldstone.¹⁸ An analysis of this approach for neon has led Barr and Davidson¹⁹ to conclude that stopping at the two-electron cluster level is insufficient. The results of Nesbet were found to be fortuitous, because errors of basis set incompleteness canceled against the overestimation of electron correlation energy that occurred at this level. Barr and Davidson further questioned the convergence of the entire procedure. More recently, at the molecular level, Dovill and Rubio [20] tested the Bethe–Goldstone (BG) approach using ammonia. Using a canonical molecular orbital basis and carrying out calculations out to the four-particle cluster level they found discouragingly slow rates of convergence. In contrast, we found, at least at the semiempirical level¹⁵ using a localized bond/antibond non-Hartree–Fock orbital basis, that the total energy is well converged. Other investigators,^{16,17,21,22} using localized basis sets of orbitals at the post-Hartree–Fock level also have had success. A formal derivation both for the wave function as well as the energy within this model can be found in the work of Roeggen who called the BG approach the “extended geminal method.”¹⁶ Roeggen and coworkers²¹ carried out applications principally on intermolecular complexes and weakly bound van der Waals dimers and trimers. A similar procedure, the “method of local increments,” has also

been used extensively by the Stollhoff, Stoll, and Fulde groups¹⁷ in their studies of solid-state properties of carbon allotropes, semiconductors, and ionic crystals.²² In contrast to these variants of the Bethe–Goldstone equations, whose focus has been mainly on the evaluation of the dynamic electron correlation energy, we employ the BG method to determine the nondynamic electron correlation energy. In this initial study we apply our algorithm to closed-shell molecules taking all possible combinations of one, two, three, and four electron pairs into account. For N_p localized electron pairs found within the molecule the resulting method then formally scales as N_p^4 once the integral transformation step into the GVB-PP basis has been performed. For this exploratory study, full CI programs using Slater determinants and the Cooper–Nesbet algorithm,²³ which treat up to eight electrons in eight active orbitals, were written. Using equations (4a)–(4f), the approximate total full CI energy was then constructed. Because, physically, the BG approach and the coupled cluster ansatz are similar approximations, at least within a localized orbital basis, we gave our method the more descriptive name *localized cluster CI*.

LCCI Tests

For an active space of four bonds and four antibonds, the LCCI method carried out to fourth

order, LCCI(4), will predict exactly the full CI energy computed within the same space. The first nontrivial test therefore occurs in calculations of molecules containing five covalent bonds or lone pairs. In Tables I–III we therefore present results for the triple-bonded acetylene and its isoelectronic analog iminoborane, HBNH. To assess the limits in predicting the nondynamic electron correlation energy, all bonds were dissociated simultaneously, separately or in some intermediate combination. For acetylene, distorted geometries were also produced by rotating the end hydrogens into a trans conformation. Because the focus here is on the convergence of the LCCI algorithm to the exact nondynamic electron correlation energy for a given active space of orbitals rather than the particular optimized set of orbitals comprising the active space, these test calculations were economically performed at the minimal basis set STO-3G and semiempirical AM1 levels. In addition to a substantially reduced integral transformation step, the active space spans the entire basis set, excluding the inner-shell atomic orbitals.²⁴ This allows one to take advantage of the rotational invariance of the exact energy and use symmetry-adapted molecular orbitals to reduce the task of the computationally intensive full CI calculations. From comparative calculations performed at the 6-311G(d,p) level we found that the minimal STO-3G basis results overestimate the nondynamic electron correlation and actually provide a more stringent test for the

TABLE I.
STO-3G Acetylene Results.

R_{CC}^a	R_{CH}^a	ω_{CCH}^b	E_{HF}^c	ϵ_0^c	ϵ_1^c	ϵ_2^c	ϵ_3^c	ϵ_4^c	CCSD(T) ^c
1.0	1.0	0.0	169.489	169.541	74.654	6.533	−0.693	−0.020	1.802
1.0	1.5	0.0	234.333	235.245	84.203	10.228	−1.501	0.076	1.835
1.0	2.0	0.0	366.039	375.970	97.162	15.670	−5.035	−0.073	−10.206
1.5	1.0	0.0	336.112	371.601	114.995	14.365	1.008	−0.175	−9.228
1.5	1.5	0.0	430.566	433.072	136.690	32.222	−1.899	−0.879	−13.509
1.5	2.0	0.0	496.674	575.769	130.115	90.576	33.484	−6.221	−246.622
2.0	1.0	0.0	648.437	651.487	168.236	18.795	0.679	−0.199	−185.546
2.0	1.0	0.0	286.123	286.258	14.773	4.151	0.248	−0.827	(d)
2.0	1.5	0.0	348.329	425.930	17.648	3.626	0.172	−0.183	−125.423
2.0	2.0	0.0	484.381	577.053	38.319	6.284	−0.010	−0.229	−107.557
1.0	1.0	45.0	182.402	184.037	71.079	7.021	−0.909	−0.113	2.124
1.5	1.5	45.0	340.240	356.651	102.564	25.728	1.337	−0.706	7.413
2.0	2.0	45.0	481.292	444.825	33.509	4.201	1.619	0.844	−61.660
1.0	1.0	90.0	210.745	217.095	82.312	9.826	−1.923	−0.093	3.008
1.5	1.5	90.0	345.004	363.102	103.984	30.846	6.340	0.438	6.295
2.0	2.0	90.0	389.986	398.594	17.611	1.409	−0.025	−0.213	−11.206

^aIn units of the equilibrium bond length. ^bHydrogens are rotated into a trans configuration with the valence bond angle, ω_{CCH} , given in degrees. ^cEnergies are in millihartrees relative to the full CI result. ^dExcited state.

TABLE II.
AM1 Acetylene Results.

R_{CC}^a	R_{CH}^a	Ω_{CCH}^b	E_{HF}^c	ε_0^c	ε_1^c	ε_2^c	ε_3^c	ε_4^c
1.0	1.0	0.0	37.143	37.146	17.813	3.072	0.221	-0.051
1.0	1.5	0.0	58.971	59.312	24.049	5.221	-0.145	-0.092
1.0	2.0	0.0	128.457	134.784	57.022	14.448	-3.450	-0.149
1.5	1.0	0.0	136.722	136.729	47.126	5.489	0.118	-0.059
1.5	1.5	0.0	155.757	156.063	55.832	15.302	-1.242	-0.893
1.5	2.0	0.0	236.665	234.700	62.839	39.359	7.522	-2.544
2.0	1.0	0.0	123.505	123.520	8.742	1.781	0.489	-0.019
2.0	1.5	0.0	150.711	151.225	10.702	1.860	0.180	0.036
2.0	2.0	0.0	229.591	234.255	22.553	2.912	-0.027	0.304
1.0	1.0	45.0	45.275	45.658	17.666	3.706	0.450	-0.025
1.5	1.5	45.0	128.939	132.168	45.108	12185	0.523	-0.150
2.0	2.0	45.0	225.669	340.495	18.550	1.008	-0.227	-0.486
1.0	1.0	90.0	60.538	61.297	22.842	6.011	1.029	0.142
1.5	1.5	90.0	139.327	142.699	52.401	17.311	2.865	0.105
2.0	2.0	90.0	175.741	336.491	11.010	0.874	0.095	0.008

^aIn units of the equilibrium bond length. ^bHydrogens are rotated into a trans configuration with the valence bond angle, ω_{CCH} , given in degrees. ^cEnergies are in millihartrees relative to the full CI result.

TABLE III.
AM1 Iminoborane Results.

R_{BN}^a	R_{XH}^b	E_{HF}^c	ε_0^c	ε_1^c	ε_2^c	ε_3^c	ε_4^c
1.0	1.0	21.213	21.303	12.131	3.005	0.302	-0.039
1.0	1.5	42.827	43.235	17.256	4.608	0.091	-0.078
1.0	2.0	114.013	160.117	48.425	24.772	-2.790	0.753
1.5	1.0	58.355	60.517	37.596	-0.710	-0.877	-0.006
1.5	1.5	167.544	170.038	134.023	-59.366	-18.521	-0.135
1.5	2.0	219.456	312.390	25.571	4.254	0.179	-0.024
2.0	1.0	43.186	43.201	24.516	0.846	0.256	0.011
2.0	1.5	61.338	61.510	31.777	9.134	0.655	-0.921
2.0	2.0	167.905	169.347	109.085	79.785	46.437	-5.413

^aIn units of the equilibrium bond length. ^bX is either boron or nitrogen; that is, the BH and NH bonds are stretched in units of their respective equilibrium bond lengths. ^cEnergies are in millihartrees relative to the full CI result.

LCCI algorithm than if more extensive basis sets were used.

All energies reported within Tables I–VI are given in millihartrees relative to the full CI benchmarks. The Hartree–Fock values, E_{HF} , then become simply the electron correlation energies for the particular minimal basis set. Unlike Hartree–Fock orbitals, which are optimized to exclusively produce the lowest energy for a single reference determinant, the GVB-PP bonding orbitals are optimized along with antibonding counterparts to produce the lowest energy for a constrained coupled cluster wave function which also takes intra-bond electron correlation into account [see eq. (1)]. The resulting reference energies, ε_0 , which are

expectation values of the single determinant reference state of doubly occupied GVB-PP localized orbitals, will then lie above the Hartree–Fock energies with the gap increasing as one moves to molecular geometries further from the equilibrium one. The total molecular energies when localized electron pairs are considered separately, in doubles, triples, and quadruples, are given by ε_1 , ε_2 , ε_3 , and ε_4 , respectively. By fourth order with the exception of the $[1.5 \times R_{CC} + 2.0 \times R_{CH}]$ geometry, all LCCI(4) energies for acetylene lie within a millihartree of the corresponding full CI results. For comparison, we also report results for the coupled cluster singles, doubles with approximate triples, method CCSD(T).²⁵ The inclusion of three

TABLE IV.
STO-3G Ethylene Results.

R_{CC}^a	R_{CH}^a	Ω_{HCCH}^b	E_{HF}^c	ε_0^c	ε_1^c	ε_2^c	ε_3^c	ε_4^c	CCSD(T) ^c
1.0	1.0	0.0	157.380	157.463	51.964	8.970	-0.638	-0.031	0.493
1.0	1.5	0.0	293.581	296.570	82.404	19.279	-1.950	-0.377	1.924
1.0	2.0	0.0	557.966	605.355	176.364	36.018	-14.645	-6.615	-50.134
1.5	1.0	0.0	283.807	284.117	62.122	6.868	-0.737	0.093	1.830
1.5	1.5	0.0	408.807	413.613	99.902	19.177	-1.636	-0.550	2.712
1.5	2.0	0.0	641.994	697.850	182.218	25.307	-12.706	4.375	-80.634
2.0	1.0	0.0	455.506	457.066	81.611	8.187	-0.429	0.669	-29.902
2.0	1.5	0.0	549.328	568.383	137.300	13.035	-3.261	-0.438	-10.651
2.0	2.0	0.0	761.435	825.456	274.770	91.410	-9.966	15.428	DNC ^d
1.0	1.0	45.0	167.158	167.437	54.634	9.311	-0.342	-0.023	0.521
1.5	1.5	45.0	423.007	429.263	100.700	19.182	-1.322	-0.623	0.317
2.0	2.0	45.0	767.816	833.433	276.026	91.849	-9.080	16.809	DNC ^d
1.0	1.0	90.0	246.996	263.529	60.422	13.498	0.270	-0.198	180.672
1.5	1.5	90.0	471.977	488.509	101.425	19.164	-1.289	-0.837	183.271
2.0	2.0	90.0	681.902	853.183	277.429	91.347	-9.097	18.952	DNC ^d

^aIn units of the equilibrium bond length. ^bThe dihedral angle, Ω_{HCCH} , is in degrees. ^cEnergies are in millihartrees relative to the full CI result. ^dDid not converge.

TABLE V.
AM1 Ethylene Results.

R_{CC}^a	R_{CH}^a	Ω_{HCCH}^b	E_{HF}^c	ε_0^c	ε_1^c	ε_2^c	ε_3^c	ε_4^c
1.0	1.0	0.0	30.367	30.370	12.685	5.059	0.402	-0.080
1.0	1.5	0.0	83.170	83.647	28.340	9.740	0.403	-0.164
1.0	2.0	0.0	235.720	250.160	107.440	14.310	-8.619	-2.100
1.5	1.0	0.0	93.586	93.592	23.455	3.806	-0.156	-0.009
1.5	1.5	0.0	142.967	143.368	43.607	13.644	1.225	-0.056
1.5	2.0	0.0	285.528	540.523	46.562	31.397	17.322	-6.550
2.0	1.0	0.0	198.608	198.716	40.069	4.897	-0.417	0.008
2.0	1.5	0.0	232.951	234.551	64.987	11.623	-1.150	-0.092
2.0	2.0	0.0	376.105	672.875	39.551	7.432	0.026	0.014
1.0	1.0	45.0	239.852	34.149	14.352	5.196	0.549	-0.052
1.5	1.5	45.0	188.975	150.976	44.244	13.355	1.427	-0.009
2.0	2.0	45.0	256.558	507.591	58.286	27.043	-1.565	-1.029
1.0	1.0	90.0	69.079	75.677	21.530	5.158	0.886	-0.068
1.5	1.5	90.0	175.417	179.844	45.060	12.830	1.572	0.021
2.0	2.0	90.0	376.654	515.573	58.714	28.771	-1.524	-1.876

^aIn units of the equilibrium bond length. ^bIn degrees. ^cEnergies are in millihartrees relative to the full CI result.

body terms at the coupled cluster level usually provides energies of millihartree accuracy compared to the full CI, even for cases when bonds are being dissociated.²⁶ We find from Table I, however, and also from Table VI, that the effects of quasidegeneracies that result when bonds are broken make the single reference CCSD(T) energies unreliable, at least at this minimal basis set level. These quasidegeneracies also make it difficult to always guarantee that the lowest lying state is

always selected when the closed-shell singlet states dissociates. To illustrate this point, in Table I we present results for two states at geometry $[2.0 \times R_{CC} + 1.0 \times R_{CH}]$. The first state, at -75.646190 hartrees, is actually the lowest lying state found from the full CI calculation. However, at the Hartree-Fock level, it lies above the next excited singlet state, which has an energy of -75.574207 hartrees. For both cases, the LCCI(4) method successfully converges to the appropriate full CI

eigenvalue. To try to ensure that the LCCI, Hartree–Fock, and subsequent full CI results are for the same state the bonding orbitals of the GVB-PP calculation were used as the starting input guess to the Hartree–Fock calculation.

The semiempirical results given in Table II essentially mirror the *ab initio* values found in Table I. We find, however, that AM1 correlation energies are substantially smaller than the STO-3G ones, even though both are minimal basis methods. Again, with the exception of the $[1.5 \times R_{\text{CC}} + 2.0 \times R_{\text{CH}}]$ geometry, the LCCI(4) values are within a millihartree of the exact full CI results. In Table III, AM1 results are also presented for the iminoborane, HBNH. The *ab initio* work of Baird and Datta²⁷ has shown that this molecule has a triple-bond comparable in strength to acetylene. The σ -bond, however, is polarized in the B^+N^- direction, whereas the π -bonds backdonate, having polarity in the B^-N^+ direction. Overall, there is a small net dipole of 0.86 D in the B^+N^- direction. As these oppositely polarized bonds are simultaneously dissociated, the LCCI(4) method achieves millihartree accuracy with the exact full CI results, with the exception of the limiting case when all five bonds are totally dissociated.

Similar results for the six covalent bonded systems, ethylene and its valence isoelectronic analog silaethylene (CSiH_4), are reported in Tables IV–VI. Bond dissociations are constrained to symmetric stretch modes so that molecular symmetry can be used to facilitate the benchmark full CI calculations. Rotations about the central molecular axis are also performed to separately break the π -bonds. Again, millihartree agreement with the exact results is found for the LCCI(4) method, except in

the case of the simultaneous dissociation of four CH bonds in ethylene. For the weaker π -bonded and more polar silaethylene, better convergence in these cases is achieved.

From the aforementioned examples we conclude that, for intermediate cases of dissociation when bonds are stretched to 1.5 times their equilibrium lengths, the LCCI(4) algorithm predicts the nondynamic electron correlation energy for the given active space with millihartree accuracy. For the simultaneous dissociation at longer bond lengths involving two or more chemical bonds results are less reliable. However, in these cases, the lowest lying state will not normally be a closed-shell singlet.

Is GVB Superior to CASSCF?

Having established the range of accurate application of the LCCI approach we turn to the question of whether GVB-CI is superior to CASSCF. In Table VII, bond dissociations of HF, H_2O , and NH_2 are compared with literature benchmark results. All values are computed with basis sets at the double zeta + polarization levels with equilibrium geometry total energies reported in hartrees and symmetric stretch geometries given as energy increases relative to the equilibrium value in kilocalories per mole. In comparison to the Hartree–Fock, GVB-PP, and CASSCF literature results, the LCCI energies lie closest to the exact full CI results. However, when comparing LCCI directly to CASSCF there are two factors that need to be distinguished: first, the size of the active space used; and, second, is whether the active orbitals

TABLE VI.
AM1 Silaethylene Results.

$R_{\text{CSi}}^{\text{a}}$	R_{XH}^{b}	E_{HF}^{c}	ϵ_0^{c}	ϵ_1^{c}	ϵ_2^{c}	ϵ_3^{c}	ϵ_4^{c}
1.0	1.0	37.085	37.103	14.180	5.061	0.216	−0.055
1.0	1.5	104.192	104.728	26.468	8.338	0.150	−0.099
1.0	2.0	282.834	295.147	71.256	19.452	−1.986	−0.848
1.5	1.0	93.614	100.796	25.635	4.489	−0.241	−0.008
1.5	1.5	155.558	169.323	47.071	17.384	2.564	−0.010
1.5	2.0	328.776	346.421	74.777	28.201	7.415	2.196
2.0	1.0	125.137	198.976	40.528	5.103	−0.520	0.001
2.0	1.5	175.924	263.973	66.032	10.969	−1.474	0.005
2.0	2.0	428.279	484.803	66.011	9.973	−0.457	0.150

^aIn units of the equilibrium bond length. ^bX is either carbon or silicon; that is, the CH and SiH bonds are symmetrically stretched in units of their respective equilibrium bond lengths. ^cEnergies are in millihartrees relative to the full CI result.

TABLE VII.
Symmetric Bond Dissociations of HF, H₂O, and NH₂.

Molecule / Method ^a	$1.0 \times r_e$	$1.5 \times r_e$	$2.0 \times r_e$
HF ($^1\Sigma^+$)			
RHF	-100.0742	71.16	143.67
GVB-PP	-100.1030	53.35	94.80
CASSCF ^b	-100.0720	52.07	93.76
CASSCF (big) ^c	-100.1692	56.32	108.59
LCCI	-100.1588	56.94	108.03
Full CI ^b	-100.2510	56.84	106.59
H ₂ O (1A_1)			
RHF	-76.0408	150.06	287.47
GVB-PP	-76.1065	108.02	172.18
CASSCF ^d	-76.1299	110.90	181.95
CASSCF(big) ^c	-76.1647	114.30	189.93
LCCI	-76.1571	113.27	187.11
Full CI ^d	-76.2566	116.23	190.99
NH ₂ (2B_2)			
ROHF	-55.5232	116.46	241.15
GVB-PP	-55.6046	101.56	190.86
CASSCF ^b	-55.6208	76.94	131.58
CASSCF(big) ^c	-55.6432	80.80	139.22
LCCI	-55.6387	80.94	141.07
Full CI ^b	-55.7426	86.23	148.78

^aAbsolute energies are given in hartrees for the equilibrium geometry, while the relative energy increases for the symmetric stretched configurations are reported in kilocalories per mole. ^bRef. 31. ^cThe active space is the same size as used in the LCCI method. ^dRef. 32.

are completely energy optimized (i.e., CASSCF) or partially optimized (i.e., GVB-PP). Because lone pair electrons are correlated with a bond/antibond orbital pair in the LCCI procedure the resulting active space will generally be larger than that found in a typical valence atomic orbital CASSCF calculation. Therefore, in Table VII, additional CASSCF calculations, which have the same number of active orbitals as the corresponding LCCI calculation, have been carried out. It can be seen that these latter results show an accuracy comparable to the LCCI method and suggest that the size of the active space is much more significant than whether the active orbitals are completely optimized or not.

Within a different context, Peterson et al.²⁸ have also reached a similar conclusion. For the prediction of spectroscopic constants of diatomics they found that using GVB followed by a multireference CI (MCI) of single and double (SD) excitations produced results as accurate as those obtained by the more traditional CASSCF/MCI(SD)

approach. In Table VIII spectroscopic constants computed with Dunning's correlation-consistent polarized-valence quadruple-zeta (cc-pVQZ) basis sets²⁹ are presented for the nonpolar F₂, polar HF, and triple-bonded N₂ diatomics and compared with the results of Peterson et al.²⁸ For each set of spectroscopic constants, nine fixed-point energy calculations were carried out about the equilibrium geometry and the results fitted to a sixth degree polynomial following the Dunham procedure.³⁰ Again, due to the presence of lone pair electrons, the LCCI method uses 14 active orbitals for F₂ and N₂, respectively, and 8 orbitals for HF. In comparison, the valence CASSCF calculations have smaller active spaces with 8 orbitals for F₂ and N₂ and 5 for HF. It is therefore perhaps not surprising that the LCCI results are more accurate than the CASSCF for F₂, marginally better for HF, and for N₂ the two methods are about equally accurate. Due to our limited computer resources, it was not practical to perform larger CASSCF calculation with same the number of active orbitals as the LCCI to check if indeed the size of the active space is the dominating factor governing the accuracy here.

Finally, equilibrium geometry optimizations are presented for small polyatomics at the 6-311G(d,p) level in Table IX. With the exception of hydrogen peroxide, computed LCCI bond distances lie within a picometer of the MP2 values for single bonds, whereas angles are predicted to be approximately within a degree of their MP2 values. For multiple bonds, where MP2 results are known to be less reliable,⁵ the LCCI results lie closer to the corresponding experimental values. Because geometry optimizations were carried out here numerically using fixed-point energy calculations, accuracy diminishes as the number of variables needing to be optimized increases. This may be the reason for the less than satisfactory agreement found in HOOH, and explains why the sample of molecules in Table IX is limited.

In summary, we have presented a new multi-configuration method, called "localized cluster CI" (LCCI), which uses a localized active space of GVB orbitals and a Bethe-Goldstone algorithm to predict the full CI energy within the active space. The procedure is computationally cheap and ideal for implementation on parallel processors. When compared with standard valence atomic orbital CASSCF, not only is the procedure orders of magnitude faster but it is also more accurate, which is probably due to the use of a larger active orbital space than that found in CASSCF. However, the

TABLE VIII.
Ground State Spectroscopic Constants of F₂, HF, and N₂.

Molecule / Method ^a	<i>r_e</i> (pm)	<i>ω_e</i> (cm ⁻¹)	<i>β_e</i> (cm ⁻¹)	<i>ω_e χ_e</i> (cm ⁻¹)	<i>α_e</i> (cm ⁻¹)
F ₂					
RHF	132.8	1263.0	1.007	6.5	0.0081
GVB-PP	148.1	654.8	0.809	16.2	0.0199
CASSCF	146.0	730.4	0.832	15.3	0.0176
LCCI	140.5	936.7	0.900	11.7	0.0129
CASSCF + 1 + 2	141.5	893.2	0.886	12.4	0.0136
Experiment	141.2	916.6	0.890	11.2	0.0138
HF					
RHF	90.15	4440.2	21.63	85.6	0.798
GVB-PP	91.39	4153.2	21.09	100.9	0.844
CASSCF	91.52	4125.9	21.03	97.2	0.839
LCCI	91.53	4224.8	21.02	85.5	0.755
CASSCF + 1 + 2	91.46	4179.4	21.06	88.6	0.790
Experiment	91.68	4138.3	20.96	89.9	0.798
N ₂					
RHF	106.6	2729.5	2.120	10.7	0.0137
GVB-PP	109.0	2427.2	2.022	14.9	0.0172
CASSCF	110.4	2339.5	1.976	13.7	0.0168
LCCI	110.1	2363.3	1.985	14.1	0.0169
CASSCF + 1 + 2	110.0	2351.6	1.988	14.0	0.0171
Experiment	109.8	2358.6	1.998	14.3	0.0173

^aAll calculations carried out with Dunning's cc-PVQZ basis sets. RHF, CASSCF, CASSCF + 1 + 2, and experimental results are taken from the work of Peterson et al.²⁸

TABLE IX.
Comparison of RHF, MP2, and LCCI Numerically Optimized Geometries.^a

Molecule	Coordinate ^b	RHF	MP2	LCCI	Experiment ^c
HOOH	O—O bond	138.5	144.4	147.7	145.2
	O—H bond	94.2	96.3	96.7	96.5
	O—O—H angle	102.7	99.3	98.9	100.0
	H—O—O—H dihedral	116.3	119.1	122.2	119.1
HOF	O—F bond	136.0	142.2	143.0	144.2
	O—H bond	94.6	96.6	96.3	96.6
	H—O—F angle	100.7	97.9	99.2	96.8
HCN	H—C bond	105.8	106.8	107.7	106.5
	C—N bond	112.6	117.0	115.1	115.3
HNC	H—N bond	105.8	100.0	100.4	99.4
	N—C bond	114.9	118.0	116.5	116.9
H ₂ O	O—H bond	94.1	95.8	96.2	95.9
	H—O—H angle	105.4	102.3	103.6	103.9
H ₂ CO	C—O bond	117.8	121.0	120.9	120.8
	C—H bond	109.5	110.6	112.1	111.6
	H—C—O angle	122.1	122.2	122.2	121.8
O ₃	O—O bond	119.5	128.4	128.4	127.8
	O—O—O angle	119.2	116.9	116.4	116.8
C ₂ H ₂	C—C bond	118.2	121.5	121.6	120.3
	C—H bond	105.6	106.5	107.6	106.1

^aAll calculations were carried out using 6-311G(d,p) basis sets. ^bBond distances are given in picometers and angles in degrees. ^cExperimental results taken from ref. 33.

LCCI method is not as versatile as CASSCF, and its range of applicability is designed for ground and low-lying excited-state potential energy surfaces.

References

1. Roos, B. O. *Adv Chem Phys* 1987, 69, 399.
2. Jaszinski, M.; Roos, B. O.; Widmark, P. O. *J Chem Phys* 1981, 75, 306.
3. Kraemer, W. P.; Roos, B. O. *Chem Phys* 1987, 118, 345.
4. Roos, B. O.; Andersson, K.; Fulscher, M. P.; Malmqvist, P. A.; Serrano-Andres, L.; Pierloot, K.; Merchán, M. *Adv Chem Phys* 1996, 93, 219.
5. Andersson, K.; Roos, B. O. *Int J Quant Chem* 1993, 45, 591.
6. Defrees, D. J.; Levi, B. A.; Pollack, S. K.; Hehre, W. J.; Binkley, J. S.; Pople, J. A. *J Am Chem Soc* 1979, 101, 4085.
7. Ladner, R. C. PhD thesis; California Institute of Technology, 1971.
8. Harding, L. B.; Goddard, W. A., III. *J Am Chem Soc* 1975, 97, 6293.
9. Carter, E. A.; Goddard, W. A., III. *J Chem Phys* 1987, 88, 1752.
10. Bobrowicz, F. W.; Goddard, W. A., III. In: Schaefer, H. F., III, *Methods of Electronic Structure Theory*; Plenum: New York, 1977.
11. Hurley, A. C.; Lennard-Jones, J. E.; Pople, J. A. *Proc R Soc* 1953, A220, 446.
12. Cullen, J. *Chem Phys* 1995, 202, 217.
13. Fischer, T. H.; Almlöf, J. *J Phys Chem* 1992, 96, 9768.
14. Nesbet, R. K. *Phys Rev.* 1966, 155, 51.
15. Cullen, J. *J Comput Chem* 1992, 13, 901.
16. Roeggen, I. *J Chem Phys* 1983, 79, 5520.
17. Stollhoff, G.; Fulde, P. *Z Phys* 1977, B26, 257; Stollhoff, G.; Fulde, P., Eds. *J Chem Phys* 1980, 73, 4548; Kiel, B.; Stollhoff, G.; Wiegand, C.; Fulde, P.; Stoll, H. *Z Phys* 1982, B46, 1.
18. Bethe, H. A.; Goldstone, J. *Proc R Soc (Lond)* 1957, A238, 551.
19. Barr, T. L.; Davidson, E. R. *Phys Rev* 1970, A1, 644.
20. Povill, A.; Rubio, J. *J Quantum Chem* 1997, 61, 35.
21. Roeggen, I.; Almlöf, J.; Ahmadi, G. R.; Wind, P. A. *J Chem Phys* 1995, 102, 7088; Roeggen, I.; Almlöf, J. *Int J Quantum Chem* 1996, 60, 453; Roeggen, I.; Wind, P. *Chem Phys* 1992, 167, 247; Roeggen, I. *J Chem Phys* 1986, 85, 262.
22. Stollhoff, G.; Bohnen, K. P. *Phys Rev* 1988, B37, 4678; Pardon, R.; Grafenstein, J.; Stollhoff, G. *Phys Rev* 1995, B51, 10556; Paulus, B.; Fulde, P.; Stoll, H. *Phys Rev* 1995, B52, 4842; Doll, K.; Dolg, M.; Stoll, H. *Phys Rev* 1996, B54, 13529.
23. Nesbet, R. K. *J Chem Phys* 1965, 43, 311.
24. For *ab initio* calculations the LCCI method optimizes inner shell atomic orbitals as part of the core in a GVB calculation. In a full CI benchmark calculation, inner shells are optimized at the Hartree-Fock level. We assume that the energy contributions of the inner shells will normally be approximately equivalent in these two cases.
25. Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J Chem Phys* 1987, 87, 5968.
26. Noga, J.; Bartlett, R. J. *J Chem Phys* 1987, 86, 7041.
27. Baird, N. C.; Datta, R. K. *Inorg Chem* 1972, 11, 17.
28. Peterson, K. A.; Kendall, R. A.; Dunning, T. H., Jr. *J Chem Phys* 1993, 99, 1930; Peterson, K. A.; Kendall, R. A.; Dunning, T. H., Jr. 1993, 99, 9790.
29. Dunning, T. H., Jr. *J Chem Phys* 1989, 90, 1007; Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J Chem Phys* 1992, 96, 6796; Woon, D. E.; Dunning, T. H., Jr. *J Chem Phys* 1993, 98, 1358.
30. Dunham, J. L. *Phys Rev* 1932, 41, 721.
31. Bauschlicher, C. W., Jr.; Taylor, P. R. *J Chem Phys* 1987, 86, 1420.
32. Bauschlicher, C. W., Jr.; Taylor, P. R. *J Chem Phys* 1987, 86, 2779.
33. Hehre, W. J.; Radom, L.; Schleyer, P. v. R. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1985.